

УДК 004.93'12

Анализ влияния модуля имитации сетчатки на возможность распознавания образов различными архитектурами нейронных сетей

Меркелов Михаил Валерьевич

merkelovmv@bmstu.ru

SPIN-код: 1900-0253

Локтев Даниил Алексеевич

loktevdan@bmstu.ru

SPIN-код: 4883-9877

МГТУ им. Н.Э. Баумана, Москва, Россия

Аннотация. В работе рассматривается влияние модуля имитации сетчатки на возможность распознавания образов различными архитектурами нейронных сетей. Были подготовлены три модели нейронных сетей. Настройки каждой из реализованных архитектур были описаны в данной работе. Модели нейронных сетей обучались на обработанных модулем имитации сетчатки наборе изображений фруктов, сфотографированных с разных ракурсов. В работе представлен набор фильтров, с помощью которых обрабатывалась выборка перед каждым из экспериментов. Проведен ряд экспериментов, проверяющих изменение точности распознавания образов в зависимости от настроек слоев модуля имитации сетчатки. Представлены результаты проведенного эксперимента и сделаны выводы по ним.

Ключевые слова: распознавание образов, классификация, нейронные сети, сетчатка глаза, обучение с учителем

Analysis of the effect of the retinal simulation module on the image recognition capability of various neural network architectures

Merkelov Mikhail Valerievich

merkelovmv@bmstu.ru

SPIN-code: 1900-0253

Loktev Daniil Alekseevich

loktevdan@bmstu.ru

SPIN-code: 4883-9877

BMSTU, Moscow, Russia

Abstract. The paper examines the effect of the retinal simulation module on the ability of image recognition by various neural network architectures. Three neural network models were prepared. The settings of each of the implemented architectures were described in this paper. Neural network models were trained on a set of fruit images processed by the retinal simulation module, photographed from different angles. The paper presents a set of filters that were used to process the sample before each experiment. A number of experiments have been conducted to test the variation in pattern recognition accuracy depending on the layer settings of the retinal simulation module. The results of the experiment are presented and conclusions are drawn on them.

Keywords: pattern recognition, classification, neural networks, retina, supervised learning

Введение. Данная работа является продолжением ранее написанных статей, в которых было дано описание модуля имитации сетчатки [1, 2], подготовлено программное обеспечение, позволяющее работать с модулем [3], проведены эксперименты, показывающие влияние модуля на точность распознавания образов [4].

В ранее проведенных исследованиях эксперименты проводились исключительно с использованием полносвязной нейронной сети. Так как современные архитектуры нейронных сетей являются комбинацией более «базовых» и простых архитектур нейронных сетей, необходимо рассмотреть влияние модуля на каждую из подобных «базовых» архитектур.

Поэтому целью данной работы является оценка и сравнение влияния модуля имитации сетчатки на возможность распознавания образов полносвязной, сверточной и рекуррентной нейронной сети.

Материалы и методы. В данной работе в качестве закона расположения программных рецепторов в слое модуля имитации будут использованы фильтры, показанные на рис. 1.

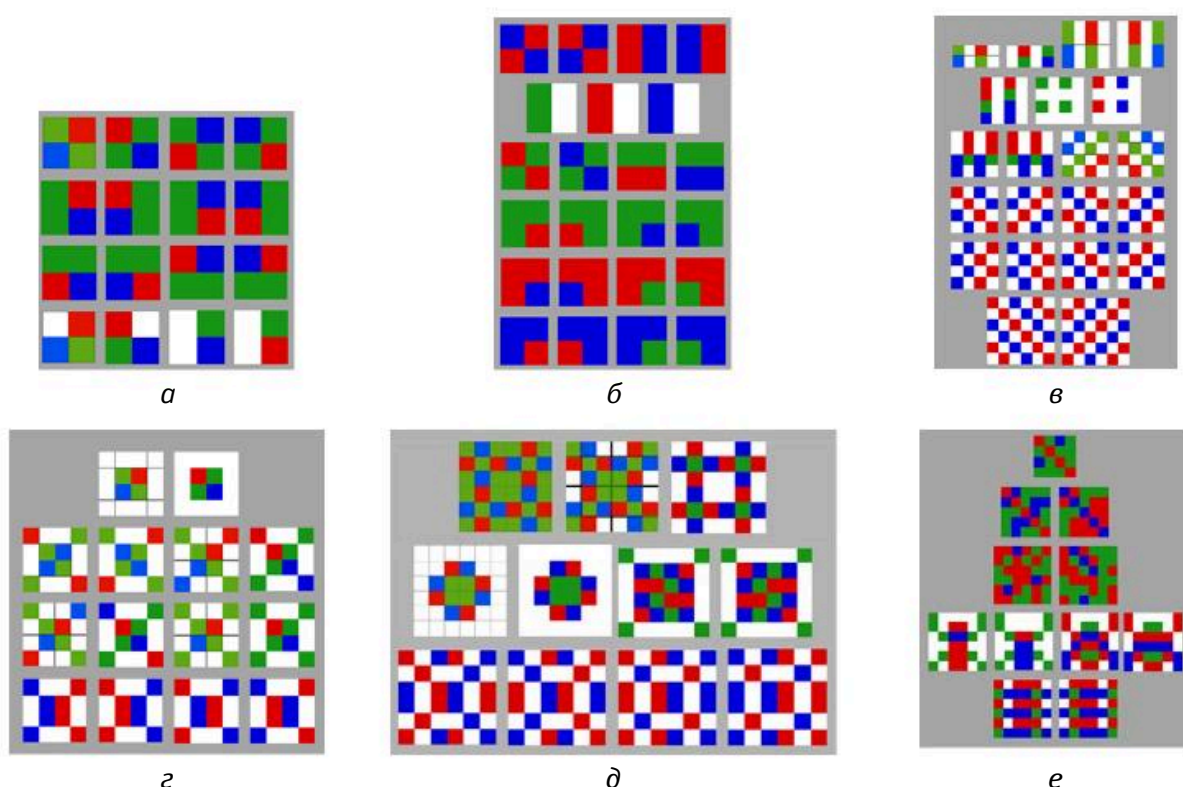


Рис. 1. Подготовленные варианты фильтров:

а — как фильтр Байера; б — двухцветные; в — линии; г — в форме «X»;
 д — как фильтр «X-Trans»; е — другое

Данные фильтры подготовлены на основе результатов прошлых экспериментов и рассмотренных ранее публикаций [4].

Настройки слоев биполярных и ганглиозных клеток совпадают с параметрами, выбранными в статье [2].

Для проверки модуля использовалась выборка [5], состоящая из фотографий 39 различных видов фруктов, сфотографированных с разных ракурсов, размером 100 на 100 пикселей. Данный набор был разделен на две части — обучающую (18505 изображений), из которой 20 % было выделено на валидационную выборку, и тестовую (6215 изображений).

На основе фильтров, показанных на рис. 1, было создано 93 варианта исходной выборки, на каждой из которых впоследствии проводился эксперимент для каждой из архитектур нейронных сетей.

Для проведения эксперимента были подготовлены на языке Python с помощью библиотеки keras три модели нейронных сетей:

- на основе полносвязной архитектуры;
- на основе сверточной архитектуры;
- на основе рекуррентной архитектуры.

Каждая из моделей содержит входной слой, принимающий изображения размером 100 на 100 пикселей, и выходной полносвязный слой из 39 нейронов с функцией активации «Softmax», необходимый для классификации 39 классов объектов выборки. Методом оптимизации при обучении был выбран стохастический градиентный спуск с моментами Нестерова со скоростью обучения 0,01. В качестве функции потерь установлена категориальная кросс-энтропия. Размер пакета составил 32 изображения.

Первая модель нейронной сети, кроме описанных ранее входного и выходного слоев, содержит еще три полносвязных слоя с функцией активации «ReLU», имитирующей реакцию реального нейрона на входной сигнал. Количество нейронов в слоях постепенно уменьшается, что может позволить нейронной сети определить характерные детали объектов классов выборки. Эта модель будет обучаться на 10 эпохах. Полная архитектура первой модели нейронной сети показана на рис. 2.

Model: "sequential"

Layer (type)	Output Shape	Param #
flatten (Flatten)	(None, 10000)	0
dense (Dense)	(None, 512)	5120512
dense_1 (Dense)	(None, 256)	131328
dense_2 (Dense)	(None, 128)	32896
dense_3 (Dense)	(None, 39)	5031
Total params: 5,289,767		
Trainable params: 5,289,767		
Non-trainable params: 0		

Рис. 2. Параметры используемой полносвязной нейронной сети

Вторая модель имитирует архитектуру VGG [6]. Основное отличие от взятой за основу архитектуры — отсутствие группы выходных полносвязных слоев. Использование нескольких последовательных сверточных слоев позволяет

извлекать более сложные признаки. Эта модель будет обучаться на 15 эпохах. Полная архитектура второй модели нейронной сети показана на рис. 3.

Model: "model"

Layer (type)	Output Shape	Param #
input_1 (InputLayer)	[(None, 100, 100, 1)]	0
conv2d_7 (Conv2D)	(None, 96, 96, 96)	2496
conv2d_8 (Conv2D)	(None, 96, 96, 96)	9312
max_pooling2d_2 (MaxPooling 2D)	(None, 47, 47, 96)	0
conv2d_9 (Conv2D)	(None, 43, 43, 192)	460992
conv2d_10 (Conv2D)	(None, 43, 43, 192)	37056
max_pooling2d_3 (MaxPooling 2D)	(None, 21, 21, 192)	0
conv2d_11 (Conv2D)	(None, 19, 19, 192)	331968
conv2d_12 (Conv2D)	(None, 19, 19, 192)	37056
conv2d_13 (Conv2D)	(None, 19, 19, 39)	7527
global_average_pooling2d_1 (GlobalAveragePooling2D)	(None, 39)	0
dense_4 (Dense)	(None, 39)	1560

=====
 Total params: 887,967
 Trainable params: 887,967
 Non-trainable params: 0

Рис. 3. Параметры используемой сверточной нейронной сети

Третья модель нейронной сети, кроме входного и выходного слоев, содержит два рекуррентных слоя и один полносвязный. На первом рекуррентном слое включен параметр, возврата последовательностей (return_sequences), так как в модели используется несколько рекуррентных слоев. Эта модель будет обучаться на 15 эпохах. Полная архитектура третьей модели нейронной сети показана на рис. 4.

Model: "model_1"

Layer (type)	Output Shape	Param #
input_2 (InputLayer)	[(None, 100, 100)]	0
lstm_2 (LSTM)	(None, 100, 256)	365568
lstm_3 (LSTM)	(None, 128)	197120
dense_6 (Dense)	(None, 64)	8256
dense_7 (Dense)	(None, 39)	2535

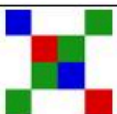
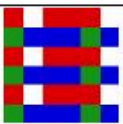
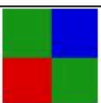
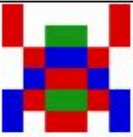
=====
 Total params: 573,479
 Trainable params: 573,479
 Non-trainable params: 0

Рис. 4. Параметры используемой рекуррентной нейронной сети

Результаты. Так как каждый запуск обучения нейронной сети производится со случайным набором весов, необходимо сделать несколько вариантов запусков для всех случаев. Поэтому для каждого фильтра на конкретной архитектуре было сделано по 10 запусков обучения. Перед каждым из запусков сначала производилась очистка всех переменных и параметров сетей, чтобы предыдущие результаты обучения не влияли на последующие запуски. Используемая выборка заново перемешивалась случайным образом.

В результате было сделано 2790 замеров точности распознавания образов (93 фильтра на каждой из трех архитектур запускалась по 10 раз). В таблице отражены лучшие и худшие результаты экспериментов.

Основные результаты проведенного эксперимента

Фильтр	Полносвязная нейронная сеть	Сверточная нейронная сеть	Рекуррентная нейронная сеть
	Средняя точность: 91 % Дисперсия: 0,000094 Ср. кв. отклонение: 0,009698	Средняя точность: 95 % Дисперсия: 0,000458 Ср. кв. отклонение: 0,021398	Средняя точность: 95 % Дисперсия: 0,000063 Ср. кв. отклонение: 0,007923
	Средняя точность: 89 % Дисперсия: 0,0004 Ср. кв. отклонение: 0,01999	Средняя точность: 85 % Дисперсия: 0,08343 Ср. кв. отклонение: 0,288842	Средняя точность: 94 % Дисперсия: 0,000092 Ср. кв. отклонение: 0,009587
	Средняя точность: 90 % Дисперсия: 0,000138 Ср. кв. отклонение: 0,011758	Средняя точность: 94 % Дисперсия: 0,000895 Ср. кв. отклонение: 0,029924	Средняя точность: 94 % Дисперсия: 0,00008 Ср. кв. отклонение: 0,008953
	Средняя точность: 76 % Дисперсия: 0,000603 Ср. кв. отклонение: 0,024555	Средняя точность: 70 % Дисперсия: 0,115028 Ср. кв. отклонение: 0,339158	Средняя точность: 78 % Дисперсия: 0,071323 Ср. кв. отклонение: 0,267064
	Средняя точность: 77 % Дисперсия: 0,000637 Ср. кв. отклонение: 0,025244	Средняя точность: 80 % Дисперсия: 0,074909 Ср. кв. отклонение: 0,273695	Средняя точность: 87 % Дисперсия: 0,000115 Ср. кв. отклонение: 0,010739
	Средняя точность: 79 % Дисперсия: 0,000537 Ср. кв. отклонение: 0,023183	Средняя точность: 78 % Дисперсия: 0,071079 Ср. кв. отклонение: 0,266607	Средняя точность: 88 % Дисперсия: 0,000169 Ср. кв. отклонение: 0,013015

Обсуждение. Из результатов экспериментов видно, что сверточная нейронная сеть во всех случаях имеет выше дисперсию и среднее квадратичное отклонение. Среди всех архитектур данная нейронная сеть показывала самое нестабильное поведение. Это может быть связано с простотой выборки, из-за чего нейронная сеть начинала переобучаться, или с реакцией на конкретные настройки модуля имитации сетчатки. Дальнейшие эксперименты планируются на более сложных выборках “Cifar” и “ImageNet”.

Также из результатов видно, что фильтр влияет на возможность распознавать образы в одном направлении сразу для всех рассмотренных архитектур, но в разных пропорциях. То есть если фильтр показал высокие результаты на полносвязной архитектуре, то он покажет высокие результаты и на других архитектурах, но при этом точность может отличаться на несколько процентов.

Рекуррентная нейронная сеть всегда показывала точность распознавания выше, чем полносвязная нейронная сеть. Отражение фильтра так же, как и у полносвязной нейронной сети, приводит к изменению точности распознавания образов.

Заключение. В работе проанализировано влияние модуля имитации сетчатки на результат распознавания образов различными архитектурами нейронных сетей. Было подготовлено три архитектуры нейронной сети, на которых проводился эксперимент.

Было выявлено, что настройки модуля изменяют точность распознавания образов в одинаковом направлении для всех рассмотренных архитектур, но в разных пропорциях. Сверточная нейронная сеть показала самое нестабильное поведение из трех рассмотренных архитектур.

В дальнейшем планируются исследования на более сложных выборках с использованием рассмотренных в данной работе архитектур нейронных сетей для подбора наилучших параметров модуля имитации сетчатки

Список источников

- [1] Меркелов М.В., Локтев Д.А. Предварительная обработка изображений с помощью модуля имитации сетчатки. *Технологии дополненной и виртуальной реальности: I науч.-практ. конф.: сб. ст.* Москва, Изд-во МГТУ им. Н.Э. Баумана, 2025, с. 212–217.
- [2] Меркелов М.В., Локтев Д.А. Влияние модуля имитации сетчатки глаза человека на качество распознавания образов нейронной сетью. *Вестник МГТУ им. Н.Э. Баумана. Сер. Приборостроение*, 2025, № 4(153), с. 102–120.
- [3] Меркелов М.В. *Модуль имитации сетчатки глаза человека*. Свидетельство о государственной регистрации программы для ЭВМ № 2025667258 РФ, 2025.
- [4] Меркелов М.В., Локтев Д.А. Влияние настроек слоя рецепторов модуля имитации сетчатки на возможность распознавания образов нейронной сетью. *Биомедсистемы-2025. XXXVIII Всерос. науч.-техн. конф. студентов, молодых ученых и специалистов: сб. тр.* Рязань, Рязанский государственный радиотехнический университет им. В.Ф. Уткина, 2025, с. 254–259.
- [5] Oltean M. Fruits-360 dataset. URL: <https://www.kaggle.com/datasets/moltean/fruits> (accessed 18.02.2026).
- [6] Simonyan K., Zisserman A. Very deep convolutional networks for large-scale image recognition. International Conference on Learning Representationse (ICLR), San Diego, 2015. <https://doi.org/10.48550/arXiv.1409.1556>